



SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Dan, Michael D.
Kaplan, Howard A.
Maiti, Pradip K.

(ii) TITLE OF INVENTION: ANTIGEN BINDING FRAGMENTS THAT
SPECIFICALLY DETECT CANCER CELLS, NUCLEOTIDES ENCODING THE
FRAGMENTS, AND USE THEREOF FOR THE PROPHYLAXIS AND
DETECTION OF CANCERS

(iii) NUMBER OF SEQUENCES: 30

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Morrison & Foerster LLP
(B) STREET: 755 Page Mill Road
(C) CITY: Palo Alto
(D) STATE: CA
(E) COUNTRY: USA
(F) ZIP: 94304-1018

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US 09/194,164
(B) FILING DATE: 09-APRIL-1999
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: PCT/US97/08962
(B) FILING DATE: 22-MAY-1997
(C) CLASSIFICATION:

(viii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/657,449
(B) FILING DATE: 22-MAY-1996
(C) CLASSIFICATION:

(ix) ATTORNEY/AGENT INFORMATION:

(A) NAME: Glaister, Debra J.
(B) REGISTRATION NUMBER: 33,888
(C) REFERENCE/DOCKET NUMBER: 31608-20001.21

(x) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (650) 813-5600
(B) TELEFAX: (650) 494-0792

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 543 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..543

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CAAGCTATT AGGTGACACT ATAGAATACT CAAGCTATGC ATCCAACGCG TTGGGAGCTC	60
TCCCCATATGG TCGACCTGCA GGCGGCCGCA CTAGTGATTT CAAGCTTCAT CACTGAACAC	120
AGAGGACTCA CCATGGAGTT TGGGCTGAGC TGGGTTTTCC TCGTTGCTCT TTTAAGAGGT	180
ATCCAGTGTGTC AGGTGCAGCT GGTGGAGTCT GGGGGAGGCG TGGTCCAGCC TGGGAGGTCC	240
CTGAGACTCT CCTGTGCAGC CTCTGGATTC CCCTTCAGAA GCTTTGCTAT GCACTGGGTC	300
CGCCAGGCTC TAGGCAAGGG GCTGGAGTGG GTGGCAGTTA TATCATATGA TGGAAAGCACT	360
AAATACTACG CAGACTCCGT GAAGGGGCGA TTCACCATCT CCAGAGACAC TTCCAAGAAC	420
ACGGTGTATC TAAAAATGAA CAGGCTGAGA ACTGAGGACA CGGCTGTCTT TTACTTGTGC	480
GAAAGACAGA GCCTGCTGGG TGACTATGAC CACTACTACG GNTTGGACGC TTGGGGAAAG	540
GGA	543

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 179 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Gln Ala Ile Val Thr Leu Asn Thr Gln Ala Met His Pro Thr Arg Trp			
1	5	10	15
Glu Leu Ser His Met Val Asp Leu Gln Ala Ala Ala Leu Val Ile Ser			
20	25	30	

Ser Phe Ile Thr Glu His Arg Gly Leu Thr Met Glu Phe Gly Leu Ser
 35 40 45

Trp Val Phe Leu Val Ala Leu Leu Arg Gly Ile Gln Cys Gln Val Gln
 50 55 60

Leu Val Glu Ser Gly Gly Val Val Gln Pro Gly Arg Ser Leu Arg
 65 70 75 80

Leu Ser Cys Ala Ala Ser Gly Phe Pro Phe Arg Ser Phe Ala Met His
 85 90 95

Trp Val Arg Gln Ala Leu Gly Lys Gly Leu Glu Trp Val Ala Val Ile
 100 105 110

Ser Tyr Asp Gly Ser Thr Lys Tyr Tyr Ala Asp Ser Val Lys Gly Arg
 115 120 125

Phe Thr Ile Ser Arg Asp Thr Ser Lys Asn Thr Val Tyr Leu Lys Met
 130 135 140

Asn Arg Leu Arg Thr Glu Asp Thr Ala Val Phe Tyr Leu Cys Glu Arg
 145 150 155 160

Gln Ser Leu Leu Gly Asp Tyr Asp His Tyr Tyr Gly Leu Asp Ala Trp
 165 170 175

Gly Lys Gly

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 543 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

TCCCTTCCCC CAAGCGTCCA ANCCGTAGTA GTGGTCATAG TCACCCAGCA GGCTCTGTCT	60
TCGCACAAG TAAAAGACAG CCGTGTCTC AGTTCTCAGC CTGTTCATTT TTAGATACAC	120
CGTGTCTTG GAAAGTGTCTC TGGAGATGGT GAATCGCCCC TTCACGGAGT CTGCGTAGTA	180
TTTAGTGCTT CCATCATATG ATATAACTGC CACCCACTCC AGCCCCTTGC CTAGAGCCTG	240
GC GGACCCAG TGCATAGCAA AGCTTCTGAA GGGGAATCCA GAGGGCTGCAC AGGAGAGTCT	300
CAGGGACCTC CCAGGGCTGGA CCACGCCTCC CCCAGACTCC ACCAGCTGCA CCTGACACTG	360
GATA CCTT AAAAGAGCAA CGAGGAAAAC CCAGCTCAGC CCAA ACTCCA TGGTGAGTCC	420

TCTGTGTTCA GTGATGAAGC TTGAAATCAC TAGTGCAGGC GCCTGCAGGT CGACCATATG	480
GGAGAGCTCC CAACGCGTTG GATGCATAGC TTGAGTATTC TATAGTGTCA CCTAAATAGC	540
TTG	543

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 450 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..450

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CTCGAGATGG ACATGGAGTT CCAGGCGCAG CTTCTCTTCC TCCTGCTACT CTGGCTCCCA	60
GATATCACCG GAGATATTGT GTTGACGCAG TCTCCAGGCA CCCTGTCTTT GTCTCCAGGG	120
GAAAGAGCCA CCCTCTCCTG CAGGGCCAGT CAGAGTGTAA GTAGCAGCTA CTTAGCCTGG	180
TACCAGCAGA AACCTGGCCA GGCTCCCAGG CTCCTCATCT ATGGTGCATC CACCAGGGCC	240
ACTGGCATGC CAGACAGGTC CAGTGGCAGT GGGTCCGGGA CAGACTTCAC TCTCACCATC	300
AGTAGACTGG AGCCTGAAGA TTTTGCAGTG TATTACTGTC AGCAGTATGG TAGCTCACCT	360
CAGACACCTC AGATCACTTT CGGCGGAGGG ACCAAGGTGG AGATCAAACG AACTGTGGCT	420
GCACCATCTG TCTTCATCTT CCCGCCATCT	450

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 150 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Leu Glu Met Asp Met Glu Phe Gln Ala Gln Leu Leu Phe Leu Leu Leu			
1	5	10	15

Leu	Trp	Leu	Pro	Asp	Ile	Thr	Gly	Asp	Ile	Val	Leu	Thr	Gln	Ser	Pro
20									25				30		
Gly	Thr	Leu	Ser	Leu	Ser	Pro	Gly	Glu	Arg	Ala	Thr	Leu	Ser	Cys	Arg
35							40					45			
Ala	Ser	Gln	Ser	Val	Ser	Ser	Ser	Tyr	Leu	Ala	Trp	Tyr	Gln	Gln	Lys
50					55						60				
Pro	Gly	Gln	Ala	Pro	Arg	Leu	Leu	Ile	Tyr	Gly	Ala	Ser	Thr	Arg	Ala
65					70				75				80		
Thr	Gly	Met	Pro	Asp	Arg	Phe	Ser	Gly	Ser	Gly	Ser	Gly	Thr	Asp	Phe
85							90						95		
Thr	Leu	Thr	Ile	Ser	Arg	Leu	Glu	Pro	Glu	Asp	Phe	Ala	Val	Tyr	Tyr
100						105						110			
Cys	Gln	Gln	Tyr	Gly	Ser	Ser	Pro	Gln	Thr	Pro	Gln	Ile	Thr	Phe	Gly
115						120						125			
Gly	Gly	Thr	Lys	Val	Glu	Ile	Lys	Arg	Thr	Val	Ala	Ala	Pro	Ser	Val
130					135					140					
Phe	Ile	Phe	Pro	Pro	Ser										
145					150										

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 450 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

AGATGGCGGG	AAGATGAAGA	CAGATGGTGC	AGCCACAGTT	CGTTTGATCT	CCACCTTGGT	60
CCCTCCGCCG	AAAGTGATCT	GAGGTGTCTG	AGGTGAGCTA	CCATACTGCT	GACAGTAATA	120
CACTGCAAAA	TCTTCAGGCT	CCAGTCTACT	GATGGTGAGA	GTGAAGTCTG	TCCCGGACCC	180
ACTGCCACTG	AACCTGTCTG	GCATGCCAGT	GGCCCTGGTG	GATGCACCAT	AGATGAGGAG	240
CCTGGGAGCC	TGGCCAGGTT	TCTGCTGGTA	CCAGGCTAAG	TAGCTGCTAC	TAACACTCTG	300
ACTGGCCCTG	CAGGAGAGGG	TGGCTCTTTC	CCCTGGAGAC	AAAGACAGGG	TGCCTGGAGA	360
CTGCGTCAAC	ACAATATCTC	CGGTGATATC	TGGGAGCCAG	AGTAGCAGGA	GGAAGAGAAG	420
CTGCGCCTGG	AACTCCATGT	CCATCTCGAG				450

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

TATGAAGACA CCAGGCCGAT ATTGTGTTGA CGCA

34

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

TATCCGGATG CAGCCACAGT TCGTTT

26

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TATTCGGACA GGTGCAGCTG GTGGAG

26

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

TATGGATCCT GAGGAGACGG TGACCGT

27

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 60 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TATATATCCG GAGGTGGTGG ATCAGGTGGA GGTGGCTCCC AGGTGCAGCT GGTGGAGTCT

60

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 46 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

ACCTCCGGAA CCGCCACCGC CAGAGACAGA TGGTGCAGCC ACATTC

46

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 918 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: join(1..906, 913..918)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GAA	TTC	ATG	AAA	AAA	ACC	GCT	ATC	GCG	ATC	GCA	GTT	GCA	CTG	GCT	GGT	48
Glu	Phe	Met	Lys	Lys	Thr	Ala	Ile	Ala	Ile	Ala	Val	Ala	Leu	Ala	Gly	
1			5				10								15	
TTC	GCT	ACC	GTT	GCG	CAG	GCC	GAT	ATT	GTG	TTG	ACG	CAG	TCT	CCA	GGC	96
Phe	Ala	Thr	Val	Ala	Gln	Ala	Asp	Ile	Val	Leu	Thr	Gln	Ser	Pro	Gly	
20			25												30	
ACC	CTG	TCT	TTG	TCT	CCA	GGG	GAA	AGA	GCC	ACC	CTC	TCC	TGC	AGG	GCC	144
Thr	Leu	Ser	Leu	Ser	Pro	Gly	Glu	Arg	Ala	Thr	Leu	Ser	Cys	Arg	Ala	
35			40												45	
AGT	CAG	AGT	GTT	AGT	AGC	AGC	TAC	TTA	GCC	TGG	TAC	CAG	CAG	AAA	CCT	192
Ser	Gln	Ser	Val	Ser	Ser	Ser	Tyr	Leu	Ala	Trp	Tyr	Gln	Gln	Lys	Pro	
50			55												60	
GGC	CAG	GCT	CCC	AGG	CTC	CTC	ATC	TAT	GGT	GCA	TCC	ACC	AGG	GCC	ACT	240
Gly	Gln	Ala	Pro	Arg	Leu	Leu	Ile	Tyr	Gly	Ala	Ser	Thr	Arg	Ala	Thr	
65			70												80	
GGC	ATG	CCA	GAC	AGG	TTC	AGT	GGC	AGT	GGG	TCC	GGG	ACA	GAC	TTC	ACT	288
Gly	Met	Pro	Asp	Arg	Phe	Ser	Gly	Ser	Gly	Ser	Gly	Thr	Asp	Phe	Thr	
85			90												95	
CTC	ACC	ATC	AGT	AGA	CTG	GAG	CCT	GAA	GAT	TTT	GCA	GTG	TAT	TAC	TGT	336
Leu	Thr	Ile	Ser	Arg	Leu	Glu	Pro	Glu	Asp	Phe	Ala	Val	Tyr	Tyr	Cys	
100			105												110	
CAG	CAG	TAT	GGT	AGC	TCA	CCT	CAG	ACA	CCT	CAG	ATC	ACT	TTC	GGC	GGA	384
Gln	Gln	Tyr	Gly	Ser	Ser	Pro	Gln	Thr	Pro	Gln	Ile	Thr	Phe	Gly	Gly	
115			120												125	
GGG	ACC	AAG	GTG	GAG	ATC	AAA	CGA	ACT	GTG	GCT	GCA	CCA	TCT	GTC	TCT	432
Gly	Thr	Lys	Val	Glu	Ile	Lys	Arg	Thr	Val	Ala	Ala	Pro	Ser	Val	Ser	
130			135												140	
GGC	GGT	GGC	GGT	TCC	GGA	GGT	GGT	GGA	TCA	GGT	GGA	GGT	GGC	TCC	CAG	480
Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly	Ser	Gln	
145			150												160	
GTG	CAG	CTG	GTG	GAG	TCT	GGG	GGA	GGC	GTG	GTC	CAG	CCT	GGG	AGG	TCC	528
Val	Gln	Leu	Val	Glu	Ser	Gly	Gly	Gly	Val	Val	Gln	Pro	Gly	Arg	Ser	
165			170												175	
CTG	AGA	CTC	TCC	TGT	GCA	GCC	TCT	GGA	TTC	CCC	TTC	AGA	AGC	TTT	GCT	576
Leu	Arg	Leu	Ser	Cys	Ala	Ala	Ser	Gly	Phe	Pro	Phe	Arg	Ser	Phe	Ala	
180			185												190	
ATG	CAC	TGG	GTC	CGC	CAG	GCT	CTA	GGC	AAG	GGG	CTG	GAG	TGG	GTG	GCA	624
Met	His	Trp	Val	Arg	Gln	Ala	Leu	Gly	Lys	Gly	Leu	Glu	Trp	Val	Ala	
195			200												205	
GTT	ATA	TCA	TAT	GAT	GGA	AGC	ACT	AAA	TAC	TAC	GCA	GAC	TCC	GTG	AAG	672
Val	Ile	Ser	Tyr	Asp	Gly	Ser	Thr	Lys	Tyr	Tyr	Ala	Asp	Ser	Val	Lys	
210			215												220	
GGC	CGA	TTC	ACC	ATC	TCC	AGA	GAC	ACT	TCC	AAG	AAC	ACG	GTG	TAT	CTA	720

Gly Arg Phe Thr Ile Ser Arg Asp Thr Ser Lys Asn Thr Val Tyr Leu			
225	230	235	240
AAA ATG AAC AGC CTG AGA ACT GAG GAC ACG GCT GTC TAT TAC TGT GCG			768
Lys Met Asn Ser Leu Arg Thr Glu Asp Thr Ala Val Tyr Tyr Cys Ala			
245	250	255	
AGA GAT CAG AGC CTG TTG GGT GAC TAT GAC CAC TAC TAC GGT TTG GAC			816
Arg Asp Gln Ser Leu Leu Gly Asp Tyr Asp His Tyr Tyr Gly Leu Asp			
260	265	270	
GTC TGG GGC AAA GGG ACC ACG GTC ACC GTC TCC TCA GGA TCC GAA CAA			864
Val Trp Gly Lys Gly Thr Thr Val Thr Val Ser Ser Gly Ser Glu Gln			
275	280	285	
AAA CTG ATC AGC GAA GAA GAT CTG AAC CAT CAC CAT CAC CAT			906
Lys Leu Ile Ser Glu Glu Asp Leu Asn His His His His His			
290	295	300	
TAGTGAAAGCTT			918
Lys Leu			

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 304 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Glu Phe Met Lys Lys Thr Ala Ile Ala Ile Ala Val Ala Leu Ala Gly			
1	5	10	15
Phe Ala Thr Val Ala Gln Ala Asp Ile Val Leu Thr Gln Ser Pro Gly			
20	25	30	
Thr Leu Ser Leu Ser Pro Gly Glu Arg Ala Thr Leu Ser Cys Arg Ala			
35	40	45	
Ser Gln Ser Val Ser Ser Ser Tyr Leu Ala Trp Tyr Gln Gln Lys Pro			
50	55	60	
Gly Gln Ala Pro Arg Leu Leu Ile Tyr Gly Ala Ser Thr Arg Ala Thr			
65	70	75	80
Gly Met Pro Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr			
85	90	95	
Leu Thr Ile Ser Arg Leu Glu Pro Glu Asp Phe Ala Val Tyr Tyr Cys			
100	105	110	
Gln Gln Tyr Gly Ser Ser Pro Gln Thr Pro Gln Ile Thr Phe Gly Gly			

115	120	125	
Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala Pro Ser Val Ser			
130	135	140	
Gly Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gln			
145	150	155	160
Val Gln Leu Val Glu Ser Gly Gly Val Val Gln Pro Gly Arg Ser			
165	170	175	
Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Pro Phe Arg Ser Phe Ala			
180	185	190	
Met His Trp Val Arg Gln Ala Leu Gly Lys Gly Leu Glu Trp Val Ala			
195	200	205	
Val Ile Ser Tyr Asp Gly Ser Thr Lys Tyr Tyr Ala Asp Ser Val Lys			
210	215	220	
Gly Arg Phe Thr Ile Ser Arg Asp Thr Ser Lys Asn Thr Val Tyr Leu			
225	230	235	240
Lys Met Asn Ser Leu Arg Thr Glu Asp Thr Ala Val Tyr Tyr Cys Ala			
245	250	255	
Arg Asp Gln Ser Leu Leu Gly Asp Tyr Asp His Tyr Tyr Gly Leu Asp			
260	265	270	
Val Trp Gly Lys Gly Thr Thr Val Thr Val Ser Ser Gly Ser Glu Gln			
275	280	285	
Lys Leu Ile Ser Glu Glu Asp Leu Asn His His His His His Lys Leu			
290	295	300	

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 918 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

AAGCTTCAC TAATGGTGAT GGTGATGGTT CAGATCTTCT TCGCTGATCA GTTTTGTTTC	60
GGATCCTGAG GAGACGGTGA CCGTGGTCCC TTTGCCAG ACGTCCAAAC CGTAGTAGTG	120
GTCATAGTCA CCCAACAGGC TCTGATCTCT CGCACAGTAA TAGACAGCCG TGTCCCTCAGT	180
TCTCAGGCTG TTCATTTTA GATACACCGT GTTCTGGAA GTGTCTCTGG AGATGGTGAA	240

TCGGCCCTTC	ACGGAGTCTG	CGTAGTATT	AGTGCTTCCA	TCATATGATA	TAAC TGCCAC	300
CCACTCCAGC	CCCTTGCTA	GAGCCTGGCG	GACCCAGTGC	ATAGCAAAGC	TTCTGAAGGG	360
GAATCCAGAG	GCTGCACAGG	AGAGTCTCAG	GGACCTCCC	GGCTGGACCA	CGCCTCCCC	420
AGACTCCACC	AGCTGCACCT	GGGAGCCACC	TCCACCTGAT	CCACCACCTC	CGGAACCGCC	480
ACCGCCAGAG	ACAGATGGTG	CAGCCACAGT	TCGTTTGATC	TCCACCTTGG	TCCCTCCGCC	540
GAAAGTGATC	TGAGGTGTCT	GAGGTGAGCT	ACCATACTGC	TGACAGTAAT	ACACTGCAAA	600
ATCTTCAGGC	TCCAGTCTAC	TGATGGTGAG	AGTGAAGTCT	GTCCCGGACC	CACTGCCACT	660
GAACCTGTCT	GGCATGCCAG	TGGCCCTGGT	GGATGCACCA	TAGATGAGGA	GCCTGGGAGC	720
CTGGCCAGGT	TTCTGCTGGT	ACCAGGCTAA	GTAGCTGCTA	CTAACACTCT	GACTGGCCCT	780
GCAGGAGAGG	GTGCCTCTTT	CCCCTGGAGA	CAAAGACAGG	GTGCCTGGAG	ACTGCGTCAA	840
CACAATATCG	GCCTGCGCAA	CGGTAGCGAA	ACCAGCCAGT	GCAACTGCGA	TCGCGATAGC	900
GGTTTTTTTC	ATGAATTC					918

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 867 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: join(1..855, 862..867)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GAA	TTC	ATG	AAA	AAA	ACC	GCT	ATC	GCG	ATC	GCA	GTT	GCA	CTG	GCT	GGT	48
Glu	Phe	Met	Lys	Lys	Thr	Ala	Ile	Ala	Ile	Ala	Val	Ala	Leu	Ala	Gly	
1			5						10					15		
TTC	GCT	ACC	GTT	GCG	CAG	GCC	GAT	ATT	GTG	TTG	ACG	CAG	TCT	CCA	GGC	96
Phe	Ala	Thr	Val	Ala	Gln	Ala	Asp	Ile	Val	Leu	Thr	Gln	Ser	Pro	Gly	
			20						25					30		
ACC	CTG	TCT	TTG	TCT	CCA	GGG	GAA	AGA	GCC	ACC	CTC	TCC	TGC	AGG	GCC	144
Thr	Leu	Ser	Leu	Ser	Pro	Gly	Glu	Arg	Ala	Thr	Leu	Ser	Cys	Arg	Ala	
			35						40					45		
AGT	CAG	AGT	GTT	AGT	AGC	AGC	TAC	TTA	GCC	TGG	TAC	CAG	CAG	AAA	CCT	192
Ser	Gln	Ser	Val	Ser	Ser	Ser	Tyr	Leu	Ala	Trp	Tyr	Gln	Gln	Lys	Pro	
			50						55					60		

GGC CAG GCT CCC AGG CTC CTC ATC TAT GGT GCA TCC ACC AGG GCC ACT		240
Gly Gln Ala Pro Arg Leu Leu Ile Tyr Gly Ala Ser Thr Arg Ala Thr		
65 70 75 80		
GGC ATG CCA GAC AGG TTC AGT GGC AGT GGG TCC GGG ACA GAC TTC ACT		288
Gly Met Pro Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr		
85 90 95		
CTC ACC ATC AGT AGA CTG GAG CCT GAA GAT TTT GCA GTG TAT TAC TGT		336
Leu Thr Ile Ser Arg Leu Glu Pro Glu Asp Phe Ala Val Tyr Tyr Cys		
100 105 110		
CAG CAG TAT GGT AGC TCA CCT CAG ACA CCT CAG ATC ACT TTC GGC GGA		384
Gln Gln Tyr Gly Ser Ser Pro Gln Thr Pro Gln Ile Thr Phe Gly Gly		
115 120 125		
GGG ACC AAG GTG GAG ATC AAA CGA ACT GTG GCT GCA TCC GGA CAG GTG		432
Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala Ser Gly Gln Val		
130 135 140		
CAG CTG GTG GAG TCT GGG GGA GGC GTG GTC CAG CCT GGG AGG TCC CTG		480
Gln Leu Val Glu Ser Gly Gly Val Val Gln Pro Gly Arg Ser Leu		
145 150 155 160		
AGA CTC TCC TGT GCA GCC TCT GGA TTC CCC TTC AGA AGC TTT GCT ATG		528
Arg Leu Ser Cys Ala Ala Ser Gly Phe Pro Phe Arg Ser Phe Ala Met		
165 170 175		
CAC TGG GTC CGC CAG GCT CTA GGC AAG GGG CTG GAG TGG GTG GCA GTT		576
His Trp Val Arg Gln Ala Leu Gly Lys Gly Leu Glu Trp Val Ala Val		
180 185 190		
ATA TCA TAT GAT GGA AGC ACT AAA TAC TAC GCA GAC TCC GTG AAG GGC		624
Ile Ser Tyr Asp Gly Ser Thr Lys Tyr Tyr Ala Asp Ser Val Lys Gly		
195 200 205		
CGA TTC ACC ATC TCC AGA GAC ACT TCC AAG AAC ACG GTG TAT CTA AAA		672
Arg Phe Thr Ile Ser Arg Asp Thr Ser Lys Asn Thr Val Tyr Leu Lys		
210 215 220		
ATG AAC AGC CTG AGA ACT GAG GAC ACG GCT GTC TAT TAC TGT GCG AGA		720
Met Asn Ser Leu Arg Thr Glu Asp Thr Ala Val Tyr Tyr Cys Ala Arg		
225 230 235 240		
GAT CAG AGC CTG TTG GGT GAC TAT GAC CAC TAC TAC GGT TTG GAC GTC		768
Asp Gln Ser Leu Leu Gly Asp Tyr Asp His Tyr Tyr Gly Leu Asp Val		
245 250 255		
TGG GGC AAA GGG ACC ACG GTC ACC GTC TCC TCA GGA TCC GAA CAA AAA		816
Trp Gly Lys Gly Thr Thr Val Thr Val Ser Ser Gly Ser Glu Gln Lys		
260 265 270		
CTG ATC AGC GAA GAA GAT CTG AAC CAT CAC CAT CAC CAT TAGTGA AAG		864
Leu Ile Ser Glu Glu Asp Leu Asn His His His His Lys		
275 280 285		

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 287 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Glu Phe Met Lys Lys Thr Ala Ile Ala Ile Ala Val Ala Leu Ala Gly
1 5 10 15

Phe Ala Thr Val Ala Gln Ala Asp Ile Val Leu Thr Gln Ser Pro Gly
20 25 30

Thr Leu Ser Leu Ser Pro Gly Glu Arg Ala Thr Leu Ser Cys Arg Ala
35 40 45

Ser Gln Ser Val Ser Ser Tyr Leu Ala Trp Tyr Gln Gln Lys Pro
50 55 60

Gly Gln Ala Pro Arg Leu Leu Ile Tyr Gly Ala Ser Thr Arg Ala Thr
65 70 75 80

Gly Met Pro Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr
85 90 95

Leu Thr Ile Ser Arg Leu Glu Pro Glu Asp Phe Ala Val Tyr Tyr Cys
100 105 110

Gln Gln Tyr Gly Ser Ser Pro Gln Thr Pro Gln Ile Thr Phe Gly Gly
115 120 125

Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala Ser Gly Gln Val
130 135 140

Gln Leu Val Glu Ser Gly Gly Val Val Gln Pro Gly Arg Ser Leu
145 150 155 160

Arg Leu Ser Cys Ala Ala Ser Gly Phe Pro Phe Arg Ser Phe Ala Met
165 170 175

His Trp Val Arg Gln Ala Leu Gly Lys Gly Leu Glu Trp Val Ala Val
180 185 190

Ile Ser Tyr Asp Gly Ser Thr Lys Tyr Tyr Ala Asp Ser Val Lys Gly
195 200 205

Arg Phe Thr Ile Ser Arg Asp Thr Ser Lys Asn Thr Val Tyr Leu Lys
210 215 220

Met Asn Ser Leu Arg Thr Glu Asp Thr Ala Val Tyr Tyr Cys Ala Arg
225 230 235 240

Asp Gln Ser Leu Leu Gly Asp Tyr Asp His Tyr Tyr Gly Leu Asp Val
245 250 255

Trp Gly Lys Gly Thr Thr Val Thr Val Ser Ser Gly Ser Glu Gln Lys
260 265 270

Leu Ile Ser Glu Glu Asp Leu Asn His His His His Lys Leu
275 280 285

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 867 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

AAGCTTCAC TAATGGTGAT GGTGATGGTT CAGATCTTCT TCGCTGATCA GTTTTGTT	60
GGATCCTGAG GAGACGGTGA CCGTGGTCCC TTTGCCAG ACGACAAAC CGTAGTAGTG	120
GTCATAGTCA CCCAACAGGC TCTGATCTCT CGCACAGTAA TAGACAGCCG TGTCCCTCAGT	180
TCTCAGGCTG TTCATTTTA GATACACCGT GTTCTGGAA GTGTCTCTGG AGATGGTGAA	240
TCGGCCCTTC ACGGAGTCTG CGTAGTATT AGTGCTTCCA TCATATGATA TAACTGCCAC	300
CCACTCCAGC CCCTTGCCTA GAGCCTGGCG GACCCAGTGC ATAGCAAAGC TTCTGAAGGG	360
GAATCCAGAG GCTGCACAGG AGAGTCTCAG GGACCTCCCA GGCTGGACCA CGCCTCCCC	420
AGACTCCACC AGCTGCACCT GTCCGGATGC AGCCACAGTT CGTTGATCT CCACCTTGGT	480
CCCTCCGCCG AAAGTGATCT GAGGTGTCTG AGGTGAGCTA CCATACTGCT GACAGTAATA	540
CACTGCAAAA TCTTCAGGCT CCAGTCTACT GATGGTGAGA GTGAAGTCTG TCCCGGACCC	600
ACTGCCACTG AACCTGTCTG GCATGCCAGT GGCCCTGGTG GATGCACCAT AGATGAGGAG	660
CCTGGGAGCC TGGCCAGGTT TCTGCTGGTA CCAGGCTAAG TAGCTGCTAC TAACACTCTG	720
ACTGGCCCTG CAGGAGAGGG TGGCTTTTC CCCTGGAGAC AAAGACAGGG TGCCTGGAGA	780
CTGCGTCAAC ACAATATCGG CCTGCGAAC GGTAGCGAAA CCAGCCAGTG CAACTGCGAT	840
CGCGATAGCG GTTTTTCA TGAATTC	867

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:
(A) NAME/KEY: Modified Base
(B) LOCATION: 5
(D) OTHER INFORMATION: May also be Thr.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Phe His Arg Tyr Ser
5

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Phe His Arg Tyr Ser Leu Pro
5

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Phe His Arg Tyr Ser Asp Tyr
5

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids

- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Phe His Arg Tyr Ser Leu Pro
5

(2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 7 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Phe His Arg Tyr Ser Pro Thr
5

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 7 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Phe His Arg Tyr Thr Pro Gly
5

(2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 7 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Phe His Arg Tyr Ser Leu Pro

(2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 7 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Phe His Arg Tyr Ser Pro Thr

5

(2) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 7 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Phe His Arg Tyr Ser Leu Pro

5

(2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 7 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Met His Arg Tyr Thr Pro Leu

5

(2) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 5 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Phe His Lys Tyr Ser
5

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Tyr His Lys Tyr Ser
5